

# Network Synchronization with Convexity\*

Guodong Shi, Alexandre Proutiere, and Karl Henrik Johansson<sup>†</sup>

## Abstract

In this paper, we establish a few new synchronization conditions for complex networks with nonlinear and nonidentical self-dynamics with switching directed communication graphs. In light of the recent works on distributed sub-gradient methods, we impose integral convexity for the nonlinear node self-dynamics in the sense that the self-dynamics of a given node is the gradient of some concave function corresponding to that node. The node couplings are assumed to be linear but with switching directed communication graphs. Several sufficient and/or necessary conditions are established for exact or approximate synchronization over the considered complex networks. These results show when and how nonlinear node self-dynamics may cooperate with the linear diffusive coupling, which eventually leads to network synchronization conditions under relaxed connectivity requirements.

**Keywords:** Coupled oscillator, Complex networks, Synchronization, Switching graphs

## 1 Introduction

The past few decades have witnessed tremendous research interest in the emergence of collective behaviors for dynamics over complex networks [26, 41, 20, 44]. The new understanding we gained lies in that certain global network-level tasks, such as synchronization or consensus, can be achieved by local interactions under cooperative couplings of individual node dynamics [45, 44, 13]. More advanced strategies have also been developed for problems like formation, swarming, optimization, and signaling [10, 20, 25, 23, 24, 15].

Synchronization problems require the node states to asymptotically reach a common trajectory or a common value over a network. In [26], a master stability function method was proposed for the local synchronization of linearly coupled oscillators, where the dynamics of each node consists of a term of nonlinear self-dynamics and another term of local linear couplings. In [45], a thorough treatment was established for synchronization of linear diffusive couplings. When the node self-dynamics is nonlinear, it was shown that the coupling strength must dominant the influence of this self-dynamics in order for global synchronization [2, 44]. Further extensions for linearly coupled oscillators have been established under

---

\*This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, and KTH SRA TNG. A preliminary version of this work has been presented at The 51st Annual Allerton Conference on Communication, Control, and Computing, Oct. 2013 [35].

<sup>†</sup>G. Shi is with the Research School of Engineering, College of Engineering and Computer Science, The Australian National University, Canberra ACT 0200, Australia. A. Proutiere and K. H. Johansson are with the ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm 10044, Sweden. Email: guodong.shi@anu.edu.au, alepro@kth.se, kallej@ee.kth.se

more restrictions on the individual self-dynamics, e.g., passivity, symmetry, and linearity [27, 28, 17, 46]. These works mainly focused on fixed interaction graph and identical self-dynamics.

Efforts have also been made on synchronization under switching interactions, non-identical node self-dynamics, or nonlinear couplings, which turned out to be far more challenging [33, 7]. Some recent improvements to classical synchronization results include [39] and [19]. In [39], connectivity requirements were relaxed to jointly connected undirected graphs, where only the union of the switching communication graphs is assumed to be connected over certain intervals, for linear agent models. In [19], the authors provided a graph comparison perspective, based on which some new graphical conditions were obtained for synchronization conditions with nonlinear node self-dynamics but fixed communication graphs.

The difficulty in analyzing synchronization conditions comes from the nontrivial coupling between node self-dynamics and the local interactions, as well as the coupling between different node states, especially under a switching communication graph. While without self-dynamics in each node, network synchronization falls to a distributed consensus problem. For consensus seeking, it has been shown that various convergence conditions can be derived based on much relaxed connectivity conditions with even directed node interactions [40, 13, 10, 30, 3, 21, 18, 36]. On the other hand, it has also been shown that if the node self-dynamics can be properly designed, this node self-dynamics can cooperate with the consensus couplings leading to distributed solutions to certain network optimization problems [23, 24, 42, 37, 14, 9, 11], which generalized the classical incremental methods for distributed optimization [38, 22, 29].

In this paper, we try to borrow the insights from consensus-based distributed optimization methods [23, 24, 42, 37, 11], with the aim of establishing some new synchronization conditions which can partially relax the in general strong assumptions on the nonlinear node self-dynamics [45, 19]. We assume that the network nodes have non-identical nonlinear self-dynamics as gradients of some concave functions. This allows for a new class of nonlinear self-dynamics which to the best of our knowledge has not been addressed in the literature. The node couplings are linear diffusive with switching directed communication graphs. Several sufficient and/or necessary conditions are established for exact or approximate synchronization of the overall node states. These results reveal when and how nonlinear node self-dynamics may cooperate with the linear consensus coupling, which leads to synchronization conditions under much relaxed connectivity requirements to the communication graphs.

The remainder of the paper is organized as follows. In Section 2, some preliminary mathematical concepts and lemmas are introduced. In Section 3, we formulate the considered network dynamics and define the problem of interest. Section 4 presents some results on fixed graphs, and then Section 5 discusses time-varying graphs. Finally some concluding remarks are given in Section 6.

## 2 Preliminaries

In this section, we introduce some notations and provide preliminary results that will be used in the rest of the paper.

## 2.1 Directed Graphs

A directed graph (digraph)  $G = (V, E)$  consists of a finite set  $V$  of nodes and an arc set  $E$ , where an arc is an ordered pair of distinct nodes of  $V$  [12]. An element  $(i, j) \in E$  describes an arc which leaves  $i$  and enters  $j$ . A *walk* in  $G$  is an alternating sequence  $\mathcal{W} : i_1 e_1 i_2 e_2 \dots e_{m-1} i_m$  of nodes  $i_\kappa$  and arcs  $e_\kappa = (i_\kappa, i_{\kappa+1}) \in E$  for  $\kappa = 1, 2, \dots, m-1$ . A walk is called a *path* if the nodes of the walk are distinct, and a path from  $i$  to  $j$  is denoted as  $i \rightarrow j$ . A digraph  $G$  is called *undirected* when for any two nodes  $i$  and  $j$ ,  $(i, j) \in E$  if and only if  $(j, i) \in E$ ; *strongly connected* if it contains path  $i \rightarrow j$  and  $j \rightarrow i$  for every pair of nodes  $i$  and  $j$ . Ignoring the direction of the arcs, the connectivity of a undirected digraph is transformed to that of the corresponding undirected graph. A time-varying graph is defined as  $G_{\sigma(t)} = (V, E_{\sigma(t)})$  where  $\sigma : [0, +\infty) \rightarrow \mathcal{Q}$  denotes a piecewise constant function, where  $\mathcal{Q}$  is a finite set containing all possible graphs with node set  $V$ . Moreover, the joint graph of  $G_{\sigma(t)}$  in time interval  $[t_1, t_2)$  with  $t_1 < t_2 \leq +\infty$  is denoted as  $G([t_1, t_2)) = \cup_{t \in [t_1, t_2)} G(t) = (V, \cup_{t \in [t_1, t_2)} E_{\sigma(t)})$ .

## 2.2 Dini Derivatives and Limit Sets

The upper *Dini derivative* of a continuous function  $h : (a, b) \rightarrow \mathbb{R}$  ( $-\infty \leq a < b \leq \infty$ ) at  $t$  is defined as

$$D^+h(t) = \limsup_{s \rightarrow 0^+} \frac{h(t+s) - h(t)}{s}.$$

When  $h$  is continuous on  $(a, b)$ ,  $h$  is non-increasing on  $(a, b)$  if and only if  $D^+h(t) \leq 0$  for any  $t \in (a, b)$ . The next result is convenient for the calculation of the Dini derivative [6, 18].

**Lemma 1** *Let  $V_i(t, x) : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  ( $i = 1, \dots, n$ ) be  $C^1$  and  $V(t, x) = \max_{i=1, \dots, n} V_i(t, x)$ . If  $\mathcal{I}(t) = \{i \in \{1, 2, \dots, n\} : V(t, x(t)) = V_i(t, x(t))\}$  is the set of indices where the maximum is reached at  $t$ , then  $D^+V(t, x(t)) = \max_{i \in \mathcal{I}(t)} \dot{V}_i(t, x(t))$ .*

Next, consider the following autonomous system

$$\dot{x} = f(x), \tag{1}$$

where  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a continuous function. Let  $x(t)$  be a solution of (1) with initial condition  $x(t_0) = x^0$ . Then  $\Omega_0 \subset \mathbb{R}^d$  is called a *positively invariant set* of (1) if, for any  $t_0 \in \mathbb{R}$  and any  $x^0 \in \Omega_0$ , we have  $x(t) \in \Omega_0$ ,  $t \geq t_0$ , along every solution  $x(t)$  of (1).

We call  $y$  a  $\omega$ -limit point of  $x(t)$  if there exists a sequence  $\{t_k\}$  with  $\lim_{k \rightarrow \infty} t_k = \infty$  such that  $\lim_{k \rightarrow \infty} x(t_k) = y$ . The set of all  $\omega$ -limit points of  $x(t)$  is called the  $\omega$ -limit set of  $x(t)$ , and is denoted as  $\Lambda^+(x(t))$ . The following lemma is well-known [32].

**Lemma 2** *Let  $x(t)$  be a solution of (1). Then  $\Lambda^+(x(t))$  is positively invariant. Moreover, if  $x(t)$  is contained in a compact set, then  $\Lambda^+(x(t)) \neq \emptyset$ .*

## 2.3 Convex Analysis

A set  $K \subset \mathbb{R}^d$  is said to be *convex* if  $(1 - \lambda)x + \lambda y \in K$  whenever  $x \in K, y \in K$  and  $0 \leq \lambda \leq 1$ . For any set  $S \subset \mathbb{R}^d$ , the intersection of all convex sets containing  $S$  is called the *convex hull* of  $S$ , denoted by  $\text{co}(S)$ .

Let  $K$  be a closed convex subset in  $\mathbb{R}^d$  and denote  $|x|_K \doteq \inf_{y \in K} |x - y|$  as the distance between  $x \in \mathbb{R}^d$  and  $K$ , where  $|\cdot|$  is the Euclidean norm. There is a unique element  $P_K(x) \in K$  satisfying  $|x - P_K(x)| = |x|_K$  associated to any  $x \in \mathbb{R}^d$  [1]. The map  $P_K$  is called the *projector* onto  $K$ . The following lemma holds [1].

**Lemma 3** (i).  $\langle P_K(x) - x, P_K(x) - y \rangle \leq 0, \quad \forall y \in K$ .

(ii).  $|P_K(x) - P_K(y)| \leq |x - y|, x, y \in \mathbb{R}^d$ .

(iii).  $|x|_K^2$  is continuously differentiable at  $x$  with  $\nabla |x|_K^2 = 2(x - P_K(x))$ .

Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a real-valued function. We call  $f$  a convex function if for any  $x, y \in \mathbb{R}^d$  and  $0 \leq \lambda \leq 1$ , it holds that  $f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$ . The following lemma states some well-known properties for convex functions.

**Lemma 4** Let  $f : \mathbb{R}^d \rightarrow \mathbb{R} \in C^1$  be a convex function.

(i).  $f(x) \geq f(y) + \langle x - y, \nabla f(y) \rangle$ .

(ii). Any local minimum is a global minimum, i.e.,  $\arg \min f = \{z : \nabla f(z) = 0\}$ .

The following lemma is established in [34] (Lemma 13, [34])

**Lemma 5** Suppose  $K \subseteq \mathbb{R}^m$  is a convex set and let  $x_a, x_b \in \mathbb{R}^m$ .

(i)  $\langle x_a - P_K(x_a), x_b - x_a \rangle \leq |x_a|_K \cdot ||x_a|_K - |x_b|_K|$ ;

(ii) If  $|x_a|_K > |x_b|_K$ , then  $\langle x_a - P_K(x_a), x_b - x_a \rangle \leq -|x_a|_K \cdot (|x_a|_K - |x_b|_K)$ .

### 3 Problem Definition

Consider a network with node set  $V = \{1, 2, \dots, N\}$ . The node interactions are modeled by a time-varying directed graph  $G_{\sigma(t)} = (V, E_{\sigma(t)})$  with  $\sigma : [0, +\infty) \rightarrow \mathcal{Q}$  being a piecewise constant function, where  $\mathcal{Q}$  is the finite set containing all possible digraphs over node set  $V$ . We assume that there is a lower bound  $\tau_D > 0$  between two consecutive switching time instants of  $\sigma(t)$ .

A node  $j$  is said to be a *neighbor* of  $i$  at time  $t$  when there is an arc  $(j, i) \in E$ , and we let  $\mathcal{N}_i(\sigma(t))$  represent the set of agent  $i$ 's neighbors at time  $t$ . Each node holds a state  $x_i(t) \in \mathbb{R}^m$ . Let  $a_{ij}(t) > 0$  be a function marking the weight associated with arc  $(j, i)$  at time  $t$ . The nodes' dynamics are described as follows:

$$\frac{d}{dt}x_i(t) = f_i(x_i(t)) + K \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j(t) - x_i(t)), \quad i = 1, \dots, N, \quad (2)$$

where  $f_i(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is a continuous function denoting the self-dynamics of node  $i$  and  $K \geq 0$  is a given constant. Let the weighted adjacency matrix be denoted as  $A_{\sigma(t)}$  where  $[A_{\sigma(t)}]_{ij} = a_{ij}(t)$  if  $j \in \mathcal{N}_i(\sigma(t))$  and  $[A_{\sigma(t)}]_{ij} = 0$  otherwise. The weighted degree matrix is then defined as  $D_{\sigma(t)} = \text{diag}(d_1(\sigma(t)), \dots, d_N(\sigma(t)))$  with  $d_i(\sigma(t)) = \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)$ . Then  $P_{\sigma(t)} = D_{\sigma(t)} - A_{\sigma(t)}$  is the time-varying Laplacian of the network representing the coupling of the node dynamics. For the time-varying weight function  $a_{ij}(t)$ , we assume that there are  $a^* > 0$  and  $a_* > 0$  such that  $a_* \leq a_{ij}(t) \leq a^*, t \in \mathbb{R}^+$ .

For the self-dynamics  $f_i$ , we first impose the following assumption.

**[A1]** There are  $F_i : \mathbb{R}^m \rightarrow \mathbb{R}, i = 1, \dots, N$  such that  $f_i = -\nabla F_i$ , where each  $F_i$  is a  $C^1$  convex function with  $\arg \min F_i \neq \emptyset$ .

**Remark 1** *System (2) is one of the standard forms in the literature for complex network synchronization, where the first term  $f_i$  represents nonlinear node self-dynamics and the second term describes linear diffusive couplings, e.g., just to name a few [45, 2, 28, 19]. On the other hand, synchronization of networks with linear (even constant) self-dynamics but nonlinear diffusive couplings is also widely studied, e.g., the Kuramoto model [16] (see [8] for a comprehensive survey).*

**Remark 2** *To investigate the synchronization of System (2) under Assumption A1 is inspired by the recent developments of distributed optimization methods [23, 24]. Evidently System (2) is the continuous-time correspondence of the distributed sub-gradient algorithm proposed in [23] for solving*

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^N F_i(z_i) \\ & \text{subject to} && z_i \in \mathbb{R}^m, z_1 = \dots = z_N. \end{aligned} \tag{3}$$

*Continuous-time solutions to the above problem have indeed been well studied, e.g., [42, 43, 9, 11, 37], where second-order dynamics are shown to be able to derive exact solutions with fixed interaction graphs [42, 11]. The current paper is however more focused on the ability of reaching a synchronization for System (2), instead of the performance serving as a continuous-time solver of (3). In fact, clearly our results are based on weaker assumptions, e.g., the  $f_i$  are not necessarily Lipschitz and the interactions are directed, switching, and unbalanced (cf., [23, 42, 43, 11]).*

**Remark 3** *Compared to the existing work [2, 27, 28, 17, 46, 33]: Assumption A1 does not require global Lipschitz condition, nor identical dynamics for the  $f_i$ . For instance, A1 allows for the case with*

$$f_i(x) = -(x - m_i)^3$$

*with  $m_i \in \mathbb{R}$  being a constant. To the best of our knowledge, network synchronization conditions under such class of self-dynamics have not been studied in the literature.*

The initial time is set to be 0. Let  $x(t) = (x_1^T(t), \dots, x_N^T(t))^T \in \mathbb{R}^{mN}$  be the Caratheodory solution of system (2) for initial condition  $x^0 = x(0)$ . We refer to [5] regarding the existence of the Caratheodory solution for (2). We introduce the following standard synchronization definition [19].

**Definition 1** *Global synchronization of System (2) is achieved if for all  $x^0 \in \mathbb{R}^{mN}$ , we have  $\lim_{t \rightarrow +\infty} |x_i(t) - x_j(t)| = 0$  for all  $i, j = 1, \dots, N$ .*

**Remark 4** *By Assumption A1 itself there might be finite-time escape for the trajectory of System (2), i.e.,  $x(t)$  approaches infinity in a finite time interval. With suitable assumptions finite-time escape can however be excluded. We refer to the coming Lemma 6, Eq. (9), and Lemma 8, respectively, which guarantee the existence of  $x(t)$  for the entire  $[0, \infty)$  under the corresponding conditions.*

## 4 Fixed Undirected Graphs – Global Results

In this section, we consider the possibility of synchronization under fixed and undirected interaction graphs. We first establish a necessary and sufficient condition for global exact synchronization, and then a global approximate synchronization condition is established. Detailed discussions regarding the feasibility of the assumptions will be presented in the end of this section.

#### 4.1 Exact Synchronization

We make another assumption on the  $F_i$ .

**[A2]**  $\{z : f_i(z) = 0\} \neq \emptyset$  is a bounded set, and  $\langle x_i - P_{\Theta_*}(x_i), f_i(x_i) \rangle \leq 0$  for all  $x_i \in \mathbb{R}^m$  and  $i = 1, \dots, N$ , where  $\Theta_* = \text{co}(\bigcup_{i \in V} \{z : f_i(z) = 0\})$ .

Note that the inequality of Assumption A2 is not a direct consequence of convexity of the  $f_i$ . However, we will later show that convexity does lead to such inequality when the argument of each  $f_i$  is in  $\mathbb{R}$ . We present the following result.

**Theorem 1** *Assume that A1 and A2 hold. Let  $G_{\sigma(t)} \equiv G$  for some fixed, undirected, and connected graph  $G$ , and let  $a_{ij}(t) \equiv a_{ji}(t) \equiv a_{ij}$  for some  $a_{ij} > 0, i, j = 1, \dots, N$ . Then global synchronization for System (2) is achieved if and only if  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ .*

*Proof.* (Necessity) We first prove the necessity statement in Theorem 1 by a contradiction argument. Suppose global synchronization is reached under the condition that  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} = \emptyset$ . Let  $x(t)$  be a trajectory of system (2)  $\Lambda^+(x(t))$  be its  $\omega$ -limit set.

First we show that  $\Lambda^+(x(t))$  is a nonempty set. Introduce

$$\theta(x(t)) := \max_{i \in V} |x_i(t)|_{\Theta_*}^2.$$

The following lemma holds (which in fact does not rely on a fixed or undirected graph), whose proof is in Appendix A.

**Lemma 6** *Let A1 and A2 hold. Then  $\theta(x(t))$  is non-increasing along each solution of System (2).*

From the above lemma we immediately know that each trajectory  $x(t)$  is contained in a compact set. Let

$$\mathcal{M} \doteq \{x = (x_1^T \dots x_N^T)^T : x_1 = \dots = x_N; x_i \in \mathbb{R}^m, i = 1, \dots, N\} \quad (4)$$

denote the consensus manifold. Based on Lemma 2 and in view of the assumption that synchronization has been reached, we conclude that  $\Lambda^+(x(t)) \subseteq \mathcal{M} \neq \emptyset$ . Moreover,  $\Lambda^+(x(t))$  is positively invariant since (2) defines an autonomous system when the interaction graph is fixed. This is to say, any trajectory of system (2) must stay within  $\Lambda^+(x(t))$  for any initial value in  $\Lambda^+(x(t))$ .

Now we take  $y \in \Lambda^+(x(t))$ . Then we have  $y = (z_*^T \dots z_*^T)^T$  for some  $z_* \in \mathbb{R}^m$ . Suppose there exist two indices  $i_1, i_2 \in \{1, \dots, N\}$  with  $i_1 \neq i_2$  such that  $f_{i_1}(z_*) \neq f_{i_2}(z_*)$ . Consider the solution of (2) for initial time 0 and initial value  $y$ . We have  $\dot{x}_{i_1}(0) \neq \dot{x}_{i_2}(0)$ . As a result, there exists a constant  $\varepsilon > 0$  such that  $x_{i_1}(t) \neq x_{i_2}(t)$  for  $t \in (0, \varepsilon)$ . In other word, the trajectory will leave the consensus manifold  $\mathcal{M}$  for  $(0, \varepsilon)$ , and therefore will also leave the set  $\Lambda^+(x(t))$ . This contradicts the fact that  $\Lambda^+(x(t))$  is positively invariant. The necessity part of Theorem 1 has been proved.

(Sufficiency) Note that  $G$  is undirected, i.e.,  $(i, j) \in E$  if and only if  $(j, i) \in E$ , and  $a_{ij} = a_{ji}$  for all  $i$  and  $j$ . We use unordered pair  $\{i, j\}$  to denote the edge between node  $i$  and  $j$ . Denote  $\mathcal{F}(z) = \sum_{i=1}^N F_i(z)$  and  $\mathcal{F}_G(x; K) = \sum_{i=1}^N F_i(x_i) + \frac{K}{2} \sum_{\{j, i\} \in E} a_{ij} |x_j - x_i|^2$ . Denote the  $N$ 'th Cartesian product of a set  $S$  as  $S^N$ . The following lemma holds with proof given in Appendix B.

**Lemma 7** Suppose  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ . Let the communication graph  $G$  be fixed, undirected, and connected. Then  $\arg \min \mathcal{F}_G(x; K) = \left(\bigcap_{i=1}^N \{z : f_i(z) = 0\}\right)^N \cap \mathcal{M}$ .

Note that

$$K \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i) + f_i(x_i) = -\nabla_{x_i} \mathcal{F}_G(x; K). \quad (5)$$

As a result,

$$\frac{d}{dt} \mathcal{F}_G(x(t); K) = -\left|\nabla \mathcal{F}_G(x; K)\right|^2 \quad (6)$$

along each trajectory of System (2). Then by LaSalle's invariance principle we have

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), \arg \min \mathcal{F}_G(x; K)) = 0.$$

Lemma 7 further ensures

$$\lim_{t \rightarrow \infty} \text{dist}\left(x(t), \left(\bigcap_{i=1}^N \{z : f_i(z) = 0\}\right)^N \cap \mathcal{M}\right) = 0$$

if  $G$  is undirected and connected. Equivalently, global synchronization is reached and we can even predict that each limit point of  $x_i(t)$  lies in  $\bigcap_{i=1}^N \{z : f_i(z) = 0\}$  for all  $i$ .

The proof of Theorem 1 is now complete.  $\square$

**Remark 5** We see from the proof above that the construction of  $\mathcal{F}_G(x)$  is critical because the convergence argument is based on the fact that the gradient of  $\mathcal{F}_G(x)$  is consistent with the interaction graph in the sense that no additional links will be introduced in the gradient.

## 4.2 Approximate Synchronization

Theorem 1 indicates that exact synchronization is impossible unless  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  is fulfilled. In this subsection, we discuss the possibility of approximate synchronization in the absence of this nonempty interaction condition. We introduce the following definition.

**Definition 2** Global  $\epsilon$ -synchronization is achieved if for all  $x^0 \in \mathbb{R}^{mN}$ , we have

$$\limsup_{t \rightarrow +\infty} |x_i(t) - x_j(t)| \leq \epsilon, \quad i, j = 1, \dots, N. \quad (7)$$

We use the following assumption.

**[A3]** (i)  $\arg \min \mathcal{F}(z) \neq \emptyset$ ; (ii)  $\arg \min \mathcal{F}_G(x; K) \neq \emptyset$  for all  $K \geq 0$ ; (iii)  $\bigcup_{K \geq 0} \arg \min \mathcal{F}_G(x; K)$  is bounded.

For  $\epsilon$ -synchronization, we present the following result.

**Theorem 2** Assume that A1 and A3 hold. Let the interaction graph  $G_{\sigma(t)} \equiv G$  for some fixed, undirected, and connected  $G$ , and let  $a_{ij}(t) \equiv a_{ji}(t) \equiv a_{ij}$  for some  $a_{ij} > 0, i, j = 1, \dots, N$ . Then for any  $\epsilon > 0$ , there exists  $K_\epsilon > 0$  such that global  $\epsilon$ -synchronization is achieved for all  $K \geq K_\epsilon$ .



*Proof.* Let's fix  $\epsilon$ . Again, since

$$K \sum_{j \in \mathcal{N}_i} a_{ij}(x_j - x_i) + f_i(x_i) = -\nabla_{x_i} \mathcal{F}_G(x; K), \quad (8)$$

the convexity of  $\mathcal{F}_G(x; K)$  ensures that

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), \arg \min \mathcal{F}_G(x; K)) = 0. \quad (9)$$

Define  $\tilde{\mathcal{F}}(x) = \sum_{i=1}^N F_i(x_i)$ . Under Assumptions A1 and A3, we have that

$$L_0 \doteq \sup \left\{ |\nabla \tilde{\mathcal{F}}(x)| : x \in \bigcup_{K \geq 0} \arg \min \mathcal{F}_G(x; K) \right\} \quad (10)$$

is a finite number. We also define

$$D_0 \doteq \sup \left\{ |z_* - x_i| : i = 1, \dots, N, x \in \bigcup_{K \geq 0} \arg \min \mathcal{F}_G(x; K) \right\}, \quad (11)$$

where  $z_* \in \arg \min \mathcal{F}$  is an arbitrarily chosen point.

Let  $p = (p_1^T \dots p_N^T)^T \in \arg \min \mathcal{F}_G(x; K)$  with  $p_i \in \mathbb{R}^m, i = 1, \dots, N$ . Let  $P$  be the Laplacian of the graph  $G$ . Since the graph is undirected and connected, we can sort the eigenvalues of the matrix  $P \otimes I_m$  as

$$0 = \lambda_1 = \dots = \lambda_m < \lambda_{m+1} \leq \dots \leq \lambda_{mN}.$$

Let  $l_1, \dots, l_{mN}$  be the orthonormal basis of  $\mathbb{R}^{mN}$  formed by the right eigenvectors of  $P \otimes I_m$ , where  $l_1, \dots, l_m$  are eigenvectors corresponding to the zero eigenvalue. Suppose  $p = \sum_{k=1}^{mN} c_k l_k$  with  $c_k \in \mathbb{R}, k = 1, \dots, mN$ .

According to (43), we have

$$\left| K(P \otimes I_m)p \right|^2 = K^2 \left| \sum_{k=m+1}^{mN} c_k \lambda_k l_k \right|^2 = K^2 \sum_{k=m+1}^{mN} c_k^2 \lambda_k^2 \leq L_0^2, \quad (12)$$

which yields

$$\sum_{k=m+1}^{mN} c_k^2 \leq \left( \frac{L_0}{K \lambda_2^*} \right)^2, \quad (13)$$

where  $\lambda_2^* > 0$  denotes the second smallest eigenvalue of  $P$ .

Now recall that

$$\mathcal{M} \doteq \{x = (x_1^T \dots x_N^T)^T : x_1 = \dots = x_N; x_i \in \mathbb{R}^m, i = 1, \dots, N\}. \quad (14)$$

is the consensus manifold. Noticing that  $\mathcal{M} = \text{span}\{l_1, \dots, l_m\}$ , we conclude from (13) that

$$\sum_{k=m+1}^{mN} c_k^2 = \left| \sum_{k=m+1}^{mN} c_k l_k \right|^2 = |p|_{\mathcal{M}}^2 = \sum_{i=1}^N |p_i - p_{\text{ave}}|^2 \leq \left( \frac{L_0}{K \lambda_2^*} \right)^2, \quad (15)$$

where  $p_{\text{ave}} = \frac{\sum_{i=1}^N p_i}{N}$ . The last equality in (15) is due to the fact that  $\mathbf{1}_N \otimes \left( \frac{\sum_{i=1}^N p_i}{N} \right)$  is the projection of  $p$  on to  $\mathcal{M}$ . From (15),  $\sum_{i=1}^N |p_i - p_{\text{ave}}|^2$  can be sufficiently small as long as  $K$  is chosen sufficiently



large. Noticing that  $\mathcal{F}$  is a  $C^1$  function, we conclude that for any  $\varsigma > 0$ , there is  $K_1(\varsigma) > 0$  such that when  $K \geq K_1(\varsigma)$ , there hold

$$|p_i - p_{\text{ave}}| \leq \varsigma, \quad i = 1, \dots, N \quad (16)$$

and

$$|\mathcal{F}(p_i) - \mathcal{F}(p_{\text{ave}})| \leq \varsigma, \quad i = 1, \dots, N. \quad (17)$$

On the other hand, with (43), we have

$$\sum_{i=1}^N f_i(p_i) = \sum_{i=1}^N f_i(p_{\text{ave}} + \hat{p}_i) = 0, \quad (18)$$

where  $\hat{p}_i = p_i - p_{\text{ave}}$ . Now according to (16) and (18), since  $F_i \in C^1$ , for any  $\varsigma > 0$ , there is  $K_2(\varsigma) > 0$  such that when  $K \geq K_2(\varsigma)$ ,

$$\left| \sum_{i=1}^N f_i(p_{\text{ave}}) \right| \leq \frac{\varsigma}{D_0}. \quad (19)$$

This implies

$$\mathcal{F}(p_{\text{ave}}) \leq \mathcal{F}(z_*) + |z_* - p_{\text{ave}}| \times \left| \sum_{i=1}^N f_i(p_{\text{ave}}) \right| \leq \mathcal{F}(z_*) + \varsigma. \quad (20)$$

Therefore, for any  $\epsilon > 0$ , we can take  $K_0 = \max\{K_1(\epsilon/2), K_2(\epsilon/2)\}$ . Then when  $K \geq K_0$ , we have

$$|p_i - p_j| \leq \epsilon; \quad \mathcal{F}(p_i) \leq \min_z \mathcal{F}(z) + \epsilon \quad (21)$$

for all  $i$  and  $j$ . Now with (9), every limit point of system (2) is contained in the set  $\arg \min \mathcal{F}_G(x; K)$ . Noting that  $p$  is arbitrarily chosen from  $\arg \min F_G(x; K)$ ,  $\epsilon$ -synchronization is achieved as long as we choose  $K_\epsilon \geq K_0$ . This completes the proof.  $\square$

From Theorems 1 and 2, we conclude that even though without the nonempty intersection condition, it is impossible to reach exact synchronization for the considered coupled dynamics, it is still possible to find a control law that guarantees approximate synchronization with arbitrary accuracy.

### 4.3 Assumption Feasibility

This subsection discusses the feasibility of Assumptions A2 and A3.

**Proposition 1** *If  $\tilde{\mathcal{F}}(x) = \sum_{i=1}^N F_i(x_i)$  is coercive, i.e.,  $\tilde{\mathcal{F}}(x) \rightarrow \infty$  as long as  $|x| \rightarrow \infty$ , then  $\{z : f_i(z) = 0\} \neq \emptyset$  is a bounded set for all  $i = 1, \dots, N$ , and A3 holds.*

*Proof.* First of all, since  $\tilde{\mathcal{F}}(x) = \sum_{i=1}^N F_i(x_i)$  is coercive, it follows straightforwardly that  $\mathcal{F}(z) = \sum_{i=1}^N F_i(z)$  and each  $F_i(z)$  are also coercive. This implies immediately that  $\{z : f_i(z) = 0\} \neq \emptyset$  is a bounded set for all  $i = 1, \dots, N$  and A3.(i) hold.

Next, Observing that  $\frac{K}{2} \sum_{\{j,i\} \in E} a_{ij} |x_j - x_i|^2 \geq 0$  for all  $x = (x_1^T \dots x_N^T)^T \in \mathbb{R}^{mN}$  and that  $\tilde{\mathcal{F}}(x) = \sum_{i=1}^N F_i(x_i)$  is coercive, we obtain that  $\arg \min \mathcal{F}_G(x; K) \neq \emptyset$  for all  $K \geq 0$ . Thus, A3.(ii) holds.

Finally, we denote  $F_* = \min_z \mathcal{F}(z) = \mathcal{F}(z_*)$ . Since  $\sum_{i=1}^N F_i(x_i)$  is coercive, there exists a constant  $M(F_*) > 0$  such that  $\sum_{i=1}^N F_i(x_i) > F_*$  for all  $|x| > M(F_*)$ . This implies

$$\mathcal{F}_G(x; K) > \mathcal{F}_G(\mathbf{1}_N \otimes z_*; K) = F_* \quad (22)$$

for all  $|x| > M$ . That is to say, the global minimum of  $\mathcal{F}_G(x; K)$  is reached within the set  $\{|x| \leq M\}$  for all  $K > 0$ . Therefore, we have

$$\bigcup_{K \geq 0} \arg \min \mathcal{F}_G(x; K) \subseteq \{|x| \leq M(F_*)\}. \quad (23)$$

This proves A3.(iii). □

**Proposition 2** Suppose  $\{z : f_i(z) = 0\} \neq \emptyset$  is a bounded set for all  $i = 1, \dots, N$  and the node state space is  $\mathbb{R}$ , i.e.,  $m = 1$ . Then A2 and A3 hold.

*Proof.* Since each  $\{z : f_i(z) = 0\}$  is a finite interval when the node state is one dimensional, it is straightforward to verifying that  $\langle x_i - P_{\Theta_*}(x_i), f_i(x_i) \rangle \leq 0$  for all  $x_i \in \mathbb{R}$ . Thus A2 holds. We now prove A3 also holds.

(i). Let  $x_i^* \in \arg \min F_i$ . Denote  $y_* = \min\{x_1^*, \dots, x_N^*\}$ . Then for any  $i = 1, \dots, N$ , we have

$$0 \geq F_i(x_i^*) - F_i(y_*) \geq -(x_i^* - y_*)f_i(y_*) \quad (24)$$

according to inequality (i) of Lemma 4. This immediately yields  $f_i(y_*) \geq 0$  for all  $i = 1, \dots, N$ .

Thus, for any  $y < y_*$ , we have

$$\mathcal{F}(y) - \mathcal{F}(y_*) \geq (y - y_*)\nabla \mathcal{F}(y_*) = -\sum_{i=1}^N (y - y_*)f_i(y_*) \geq 0, \quad (25)$$

which implies  $\mathcal{F}(y) \geq \mathcal{F}(y_*)$  for all  $y < y_*$ .

A symmetric analysis leads to that  $\mathcal{F}(y) \geq \mathcal{F}(y^*)$  for all  $y > y^*$  with  $y^* = \max\{x_1^*, \dots, x_N^*\}$ . Therefore, we obtain  $\mathcal{F}(y) \geq \min\{\mathcal{F}(y_*), \mathcal{F}(y^*)\}$  for all  $y \neq [y_*, y^*]$ . This implies that a global minimum is reached within the interval  $[y_*, y^*] = \text{co}\{x_1^*, \dots, x_N^*\}$  and A3.(i) thus follows.

(ii). Introduce the following cube in  $\mathbb{R}^N$ :

$$\mathcal{C}_*^\eta \doteq \left\{ x = (x_1^T \dots x_N^T)^T : x_i \in [y_* - \eta, y^* + \eta], i = 1, \dots, N \right\},$$

where  $\eta > 0$  is a given constant.

*Claim.* For any  $K \geq 0$ ,  $\mathcal{C}_*^\eta$  is an invariant set of System (2).

Define  $\Psi(x(t)) = \max_{i \in \mathcal{V}} x_i(t)$ . Then based on Lemma 1, we have

$$\begin{aligned} D^+ \Psi(x(t)) &= \max_{i \in \mathcal{I}_0(t)} \frac{d}{dt} x_i(t) \\ &= \max_{i \in \mathcal{I}_0(t)} \left[ \sum_{j \in \mathcal{N}_i} a_{ij} (x_j - x_i) + f_i(x_i) \right] \\ &\leq \max_{i \in \mathcal{I}_0(t)} [f_i(x_i)], \end{aligned} \quad (26)$$

where  $\mathcal{I}_0(t)$  denotes the index set which contains all the nodes reaching the maximum for  $\Psi(x(t))$ .

Since

$$0 \geq F_i(x_i^*) - F_i(y_* + \eta) \geq -(x_i^* - y_* - \eta)f_i(y_* + \eta), \quad i = 1, \dots, N \quad (27)$$

we have  $f_i(y_* + \eta) \leq 0$  for all  $i = 1, \dots, N$ . As a result, we obtain

$$D^+\Psi(x(t)) \Big|_{\Psi(x(t))=y_*+\eta} \leq 0, \quad (28)$$

which implies  $\Psi(x(t)) \leq y_* + \eta$  for all  $t \geq t_0$  under initial condition  $\Psi(x(t_0)) \leq y_* + \eta$ . Similar analysis ensures that  $\min_{i \in V} x_i(t) \geq y_* - \eta$  for all  $t \geq t_0$  as long as  $\min_{i \in V} x_i(t_0) \geq y_* - \eta$ . This proves the claim.

Note that every trajectory of system (2) asymptotically reaches  $\arg \min \mathcal{F}_G(x; K)$ . This immediately leads to that  $\mathcal{F}_G(x; K)$  reaches its minimum within  $\mathcal{C}_*^\eta$  for any  $K \geq 0$  since  $\mathcal{C}_*^\eta$  is an invariant set. Then A3.(ii) holds.

(iii). Since  $\arg \min F_i$  is bounded for  $i = 1, \dots, N$ , there exist  $b_i \leq d_i, i = 1, \dots, N$  such that  $\arg \min F_i = [b_i, d_i]$ . Define  $b_* = \min\{b_1, \dots, b_N\}$  and  $d^* = \max\{d_1, \dots, d_N\}$ . We will prove the conclusion by showing  $\arg \min \mathcal{F}_G(x; K) \subseteq \mathcal{C}_*$  for all  $K \geq 0$ , where

$$\mathcal{C}_* \doteq \left\{ x = (x_1^T \dots x_N^T)^T : x_i \in [b_*, d^*], i = 1, \dots, N \right\}.$$

Let  $z = (z_1, \dots, z_N)^T \in \arg \min \mathcal{F}_G(x; K)$ . First we show  $\max\{z_1, \dots, z_N\} \leq d^*$  by a contradiction argument. Suppose  $\max\{z_1, \dots, z_N\} > d^*$ .

Now let  $i_1, \dots, i_k$  be the nodes reaching the maximum state, i.e.,  $z_{i_1} = \dots = z_{i_k} = \max\{z_1, \dots, z_N\}$ . There will be two cases.

- Let  $k = N$ . We have  $z_1 = \dots = z_N = y$  in this case. Then for all  $i$  and  $x_i^* \in \arg \min F_i$ , we have

$$0 > F_i(x_i^*) - F_i(y) \geq -(x_i^* - y)f_i(y) \quad (29)$$

which yields  $f_i(y) > 0, i = 1, \dots, N$  since  $y > d^*$ . This immediately leads to

$$\mathcal{F}_G(z; K) = \mathcal{F}(y) > \min \mathcal{F} \geq \min \mathcal{F}_G(z; K), \quad (30)$$

which contradicts the fact that  $z \in \arg \min \mathcal{F}_G(x; K)$ .

- Let  $k < N$ . Then we denote  $s_* = \max\{z_i : i \notin \{i_1, \dots, i_k\}, i = 1, \dots, N\}$ , which is actually the second largest value in  $\{z_1, \dots, z_N\}$ . We define a new point  $\hat{z} = (\hat{z}_1, \dots, \hat{z}_N)^T$  by  $\hat{z}_i = z_i, i \notin \{i_1, \dots, i_k\}$  and

$$\hat{z}_i = \begin{cases} d^*, & \text{if } s_* < d^* \\ s_*, & \text{otherwise} \end{cases} \quad (31)$$

for  $i \in \{i_1, \dots, i_k\}$ . Then it is easy to obtain that  $\mathcal{F}_G(z; K) > \mathcal{F}_G(\hat{z}; K)$ , which again contradicts the choice of  $z$ .

Therefore, we have proved that  $\max\{z_1, \dots, z_N\} \leq d^*$ . Based on a symmetric analysis we also have  $\min\{z_1, \dots, z_N\} \geq b_*$ . Therefore, we obtain  $\arg \min \mathcal{F}_G(x; K) \subseteq \mathcal{C}_*$  for all  $K \geq 0$  and A3.(iii) follows.  $\square$

## 5 Time-varying Directed Graphs – Global Exact and Semi-global Approximate Synchronization

In this section, we consider time-varying graphs. We introduce the following definition [13, 18].

**Definition 3**  $G_{\sigma(t)}$  is said to be uniformly jointly strongly connected if there exists a constant  $T > 0$  such that  $G([t, t + T))$  is strongly connected for any  $t \geq 0$ .

We present the following result.

**Theorem 3** Let A1 hold. Suppose  $G_{\sigma(t)}$  is uniformly jointly strongly connected and  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  contains at least one interior point. Then global synchronization is achieved for System (2). In fact, for any initial value  $x^0$ , there exists  $x_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$  such that  $\lim_{t \rightarrow \infty} x_i(t) = x_*$  for all  $i \in V$ .

Note that the condition  $\lim_{t \rightarrow \infty} x_i(t) = x_*$  is indeed a stronger conclusion than our definition of synchronization as Theorem 3 guarantees that all the node states converge to a common point. We will see from the proof of Theorem 3 that this state convergence highly relies on the existence of an interior point of  $\bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . In the absence of such an interior point condition, it turns out that global synchronization still stands. We present another theorem stating the fact.

**Theorem 4** Let A1 hold. Suppose  $G_{\sigma(t)}$  is uniformly jointly strongly connected and  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ . Then global synchronization is achieved for System (2).

For  $\epsilon$ -synchronization under switching interactions, we present the following result.

**Theorem 5** Let A1 and A2 hold. Suppose  $G_{\sigma(t)}$  is uniformly jointly strongly connected. Then for any  $\epsilon > 0$  and any initial value  $x^0 \in \mathbb{R}^{mN}$ , there exist a sufficiently small  $T_\epsilon^\dagger(x^0) > 0$  and a sufficiently large  $K_\epsilon^\dagger(x^0)$  such that  $\epsilon$ -synchronization is achieved under  $x^0$  for all  $T \leq T_\epsilon^\dagger(x^0)$  and  $K \geq K_\epsilon^\dagger(x^0)$ .

Note that compared to the results under discrete-time dynamics [23, 24], Theorems 3 and 4 stand on quite general assumptions, which applies to the case when the  $\{z : f_i(z) = 0\}$  are unbounded. Compared to Theorem 2, Theorem 5 is semi-global in the sense that the control gain  $K_\epsilon^\dagger(x^0)$  depends on the initial value. With switching interaction graphs, it becomes fundamentally difficult to characterize the limit set of the trajectories, and a general global result as Theorem 2 may not hold.

The remaining of this section presents the proofs of the above results. We first present some useful lemmas, and then the proofs of Theorems 3, 4, and 5 will be established, respectively.

### 5.1 Preliminary Lemmas

We establish three useful lemmas in this subsection. Suppose  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  and take  $z_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . We define

$$V_i(t) = |x_i(t) - z_*|^2, \quad i = 1, \dots, N, \quad (32)$$

and

$$V(t) = \max_{i=1,\dots,N} V_i(t). \quad (33)$$

The following lemma holds, whose proof is given in Appendix C.

**Lemma 8** *Let A1 hold. Suppose  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ . Then along any trajectory of System (2), we have  $D^+V(t) \leq 0$  for all  $t \in \mathbb{R}^+$ .*

A direct consequence of Lemma 8 is that when  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ , we have

$$\lim_{t \rightarrow \infty} V(t) = d_*^2 \quad (34)$$

for some  $d_* \geq 0$  along any trajectory of system (2) with control law  $\mathcal{J}_*(n_i, g_i)$ . However, it is still unclear whether  $V_i(t)$  converges or not. We establish another lemma indicating that with proper connectivity condition for the communication graph, all  $V_i(t)$ 's have the same limit  $d_*^2$ . The following Lemma holds with proof given in Appendix D.

**Lemma 9** *Let A1 hold. Suppose  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  and  $G_{\sigma(t)}$  is uniformly jointly strongly connected. Then along any trajectory of System (2), we have  $\lim_{t \rightarrow \infty} V_i(t) = d_*^2$  for all  $i$ .*

Finally, the next lemma shows that each  $x_i(t)$  asymptotically reaches  $\arg \min F_i$  along the trajectories of system (2), whose proof is in Appendix E.

**Lemma 10** *Let A1 hold. Suppose  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  and  $G_{\sigma(t)}$  is uniformly jointly strongly connected. Then along any trajectory of system (2), we have*

$$\limsup_{t \rightarrow \infty} |x_i(t)|_{\arg \min F_i} = 0$$

for all  $i$ .

## 5.2 Proofs of Statements

### 5.2.1 Proof of Theorem 3

The proof of Theorem 3 relies on the following lemma.

**Lemma 11** *Let  $z_1, \dots, z_{m+1} \in \mathbb{R}^m$  and  $d_1, \dots, d_{m+1} \in \mathbb{R}^+$ . Suppose there exist solutions to equations (with variable  $y$ )*

$$\begin{cases} |y - z_1|^2 = d_1; \\ \vdots \\ |y - z_{m+1}|^2 = d_{m+1}. \end{cases} \quad (35)$$

*Then the solution is unique if  $\text{rank}(z_2 - z_1, \dots, z_{m+1} - z_1) = m$ .*

*Proof.* Take  $j > 1$  and let  $y$  be a solution to the equations. Noticing that

$$\langle y - z_1, y - z_1 \rangle = d_1; \quad \langle y - z_j, y - z_j \rangle = d_j$$

we obtain

$$\langle y, z_j - z_1 \rangle = \frac{1}{2} \left( d_1 - d_j + |z_j|^2 - |z_1|^2 \right), \quad j = 2, \dots, m+1. \quad (36)$$

The desired conclusion follows immediately.  $\square$

Let  $r_* = (r_1^T \dots r_N^T)^T$  be a limit point of a trajectory of System (2). Based on Lemma 9, we have  $\lim_{t \rightarrow \infty} V_i(t) = d_*$  for all  $z_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . This is to say,  $|r_i - z_*| = d_*$  for all  $i$  and  $z_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . Since  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$  contains at least one interior point, it is obvious to see that we can find  $z_1, \dots, z_{m+1} \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$  with  $\text{rank}(z_2 - z_1, \dots, z_{m+1} - z_1) = m$  and  $d_1, \dots, d_{m+1} \in \mathbb{R}^+$ , such that each  $r_i, i = 1, \dots, N$  is a solution of equations (35). Then based on Lemma 11, we conclude that  $r_1 = \dots = r_N$ . Next, with Lemma 10, we have  $|r_i|_{\arg \min F_i} = 0$ . This implies that  $r_1 = \dots = r_N \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ , i.e., global synchronization is achieved.

We turn to state convergence. We only need to show that  $r_*$  is unique along any trajectory of System (2). Now suppose  $r_*^1 = \mathbf{1}_N \otimes r^1$  and  $r_*^2 = \mathbf{1}_N \otimes r^2$  are two different limit points with  $r^1 \neq r^2 \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . According to the definition of a limit point, we have that for any  $\varepsilon > 0$ , there exists a time instant  $t_\varepsilon$  such that  $|x_i(t_\varepsilon) - r^1| \leq \varepsilon$  for all  $i$ . Note that Lemma 8 indicates that the disc  $B(r^1, \varepsilon) = \{y : |y - r^1| \leq \varepsilon\}$  is an invariant set for initial time  $t_\varepsilon$ . While taking  $\varepsilon = |r^1 - r^2|/4$ , we see that  $r^2 \notin B(r^1, |r^1 - r^2|/4)$ . Thus,  $r^2$  cannot be a limit point.

Now that the limit point is unique along any trajectory of System (2), we denote it as  $\mathbf{1}_N \otimes x_*$  with  $x_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . Then we have  $\lim_{t \rightarrow \infty} x_i(t) = x_*$  for all  $i = 1, \dots, N$ . This completes the proof.

## 5.2.2 Proof of Theorem 4

In this subsection, we prove Theorem 4. We need the following lemma on robust consensus, which is a special case of Theorem 4.1 and Proposition 4.10 in [36].

**Lemma 12** *Consider the following dynamics for the considered network model:*

$$\dot{x}_i = K \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + w_i(t), \quad i \in \mathcal{V} \quad (37)$$

where  $K > 0$  is a given constant,  $a_{ij}(t)$  are weight functions satisfying our network model, and  $w_i(t)$  is a piecewise continuous function. Let  $G_{\sigma(t)}$  be uniformly jointly strongly connected with respect to  $T > 0$ .

- (i). There holds  $\lim_{t \rightarrow +\infty} |x_i(t) - x_j(t)| = 0$  for all  $i, j \in \mathcal{V}$  if  $\lim_{t \rightarrow \infty} w_i(t) = 0, i \in \mathcal{V}$ .
- (ii). For any  $\epsilon > 0$ , there exist a sufficiently small  $T_\epsilon > 0$  and sufficiently large  $K_\epsilon$  such that

$$\limsup_{t \rightarrow +\infty} |x_i(t) - x_j(t)| \leq \epsilon \|w(t)\|_\infty$$

for all initial value  $x^0$  when  $K \geq K_\epsilon$  and  $T \leq T_\epsilon$ , where  $\|w(t)\|_\infty := \max_{i \in \mathcal{V}} \sup_{t \in [0, \infty)} |w_i(t)|$ .

Lemma 10 indicates that  $\limsup_{t \rightarrow \infty} |x_i(t)|_{\arg \min F_i} = 0$  for all  $i$ , which yields

$$\lim_{t \rightarrow \infty} f_i(x_i(t)) = 0 \quad (38)$$

for all  $i$  according to Assumption A1. Then global synchronization follows immediately from Lemma 12.(i). Again by Lemma 10, we further conclude that

$$\limsup_{t \rightarrow \infty} \text{dist}(x_i(t), \bigcap_{i=1}^N \{z : f_i(z) = 0\}) = 0.$$

The desired conclusion thus follows.

### 5.2.3 Proof of Theorem 5

From Lemma 6 we know that  $\theta(x(t)) = \max_{i \in V} |x_i(t)|_{\Theta_*}^2$  is non-increasing under A2. As a result, we conclude that

$$x(t) \in \Gamma(x^0) := \left\{ z \in \mathbb{R}^{mN} : \theta(z) \leq \theta(x^0) \right\}$$

for all  $t \geq 0$ . Again by Assumption A2,  $\Gamma(x^0)$  is a compact set. We can thus define

$$\hbar(x^0) := \max_{i \in V} \sup \left\{ |f_i(z_i)| : z = (z_1 \dots z_N)^T \in \Gamma(x^0) \right\}.$$

Now along the trajectory  $x(t)$  of (2) with initial value  $x^0$ , we have

$$|f_i(x_i(t))| \leq \hbar(x^0)$$

for all  $t \geq 0$ . Then the desired  $\epsilon$ -synchronization result follows immediately from Lemma 12.(ii).

## 6 Conclusions

In light of recent works on consensus-based distributed optimization methods, we have established some conditions on the synchronization problems of coupled oscillators. We assumed that the network nodes have non-identical nonlinear self-dynamics which are gradients of some concave functions. This allowed for functions that are not passive or globally Lipschitz. The node interactions were under switching directed communication graphs. Some sufficient and/or necessary conditions are established regarding exact or approximate synchronization of the overall node states. These results revealed when and how nonlinear node self-dynamics can cooperate with the linear consensus coupling and reach synchronization with much relaxed connectivity conditions. Some interesting future generalizations include the exact convergence rate to a synchronization under strict convexity, and synchronization conditions with constrained node states.

## Acknowledgment

The authors thank Prof. Angelia Nedić for helpful discussions as well as for her pointing out relevant literature.



## Appendix

### A Proof of Lemma 6

Denote  $\mathcal{I}^\dagger(t) := \{i \in \mathcal{V} : |x_i(t)|_{\Theta_*}^2 = \theta(x(t))\}$ . Then from Lemma 1 and Lemma 3.(iii), we know

$$D^+\theta(x(t)) = 2 \max_{i \in \mathcal{I}^\dagger(t)} \left\langle x_i(t) - P_{\Theta_*}(x_i(t)), f_i(x_i(t)) + K \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j(t) - x_i(t)) \right\rangle. \quad (39)$$

Now with Lemma 5.(ii), there holds

$$\left\langle x_i(t) - P_{\Theta_*}(x_i(t)), x_j(t) - x_i(t) \right\rangle \leq 0 \quad (40)$$

for all  $i \in \mathcal{I}^\dagger(t)$  and  $j$  due to the definition of  $\mathcal{I}^\dagger(t)$  and  $\theta(x(t))$ . Combining (39), (40), and Assumption A2 we further know

$$D^+\theta(x(t)) \leq 2 \max_{i \in \mathcal{I}^\dagger(t)} \left\langle x_i(t) - P_{\Theta_*}(x_i(t)), f_i(x_i(t)) \right\rangle \leq 0. \quad (41)$$

The desired lemma thus follows.  $\square$

### B Proof of Lemma 7

When  $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$ , it is clear that  $\arg \min \mathcal{F} = \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ .

Now take  $x_* = (p_*^T \dots p_*^T)^T \in (\bigcap_{i=1}^N \{z : f_i(z) = 0\})^N \cap \mathcal{M}$ , where  $p_* \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ . First we have  $x_* \in \arg \min_x \sum_{i=1}^N F_i(x_i)$ . Second we have  $x_* \in \arg \min_x \frac{K}{2} \sum_{\{j,i\} \in \mathcal{E}} a_{ij} |x_j - x_i|^2$ . Therefore, we conclude that  $x_* \in \arg \min \mathcal{F}_G(x; K)$ . This gives

$$\arg \min \mathcal{F}_G(x; K) \supseteq \left( \bigcap_{i=1}^N \{z : f_i(z) = 0\} \right)^N \cap \mathcal{M}. \quad (42)$$

On the other hand, convexity gives

$$\arg \min \mathcal{F}_G(x; K) = \left\{ x : -K(P \otimes I_m)x = \left( (f_1(x_1))^T \dots (f_N(x_N))^T \right)^T \right\}, \quad (43)$$

where  $\otimes$  represents the Kronecker product,  $I_m$  is the identity matrix in  $\mathbb{R}^m$ , and  $P$  is the Laplacian of the graph  $G$ . Noticing that

$$(\mathbf{1}_N^T \otimes I_m)(P \otimes I_m) = \mathbf{1}_N^T P \otimes I_m = 0,$$

where  $\mathbf{1}_N = (1 \dots 1)^T \in \mathbb{R}^N$ , we have

$$\left( (\mathbf{1}_N^T \otimes I_m) \left( (f_1(x_1))^T \dots (f_N(x_N))^T \right)^T \right) = \sum_{i=1}^N f_i(x_i) = 0 \quad (44)$$

for any  $x \in \arg \min \mathcal{F}_G(x; K)$ .

Take  $x^* = (q_1^T \dots q_N^T)^T \in \arg \min \mathcal{F}_G(x; K)$ . Suppose there exist two indices  $i_*$  and  $j_*$  such that

$$f_{i_*}(q_{i_*}) \neq f_{j_*}(q_{j_*}).$$

Then at least one of  $f_{i_*}(q_{i_*})$  and  $f_{j_*}(q_{j_*})$  must be nonzero. Taking  $\hat{p} \in \bigcap_{i=1}^N \{z : f_i(z) = 0\}$ , we have

$$\sum_{i=1}^N F_i(q_i) > \sum_{i=1}^N F_i(\hat{p})$$

because for  $x = (x_1^T \dots x_N^T)^T \in \arg \min \sum_{i=1}^N F_i(x_i)$ , we have  $f_i(x_i) = 0, i = 1, \dots, N$ . Consequently, for  $w_* = (\hat{p}^T \dots \hat{p}^T)^T$ , we have

$$\mathcal{F}_G(x^*; K) > \mathcal{F}_G(w_*; K)$$

which is impossible according to the definition of  $x^*$  so that such  $i_*$  and  $j_*$  cannot exist. In light of (44), this immediately implies  $f_i(q_i) = 0, i = 1, \dots, N$ , or equivalently

$$q_i \in \{z : f_i(z) = 0\}, i = 1, \dots, N \quad (45)$$

for all  $x^* = (q_1^T \dots q_N^T)^T \in \arg \min \mathcal{F}_G(x)$ . Therefore, we conclude from (45) that  $\sum_{i=1}^N F_i(q_i) = \sum_{i=1}^N F_i(p_*)$ , and this implies

$$\sum_{\{j,i\} \in E} a_{ij} |q_j - q_i|^2 = 0$$

as long as  $x^* = (q_1^T \dots q_N^T)^T \in \arg \min \mathcal{F}_G(x)$ . The connectivity of the communication graph thus further guarantees that  $q_1 = \dots = q_N$ , so we have proved that  $x^* \in \left(\bigcap_{i=1}^N \{z : f_i(z) = 0\}\right)^N \cap \mathcal{M}$ . Consequently, we obtain

$$\arg \min \mathcal{F}_G(x; K) \subseteq \left(\bigcap_{i=1}^N \{z : f_i(z) = 0\}\right)^N \cap \mathcal{M}. \quad (46)$$

The desired lemma holds from (42) and (46).  $\square$

## C Proof of Lemma 8

Based on Lemma 1, we have

$$\begin{aligned} D^+V(t) &= \max_{i \in \mathcal{I}(t)} \frac{d}{dt} V_i(t) \\ &= \max_{i \in \mathcal{I}(t)} 2 \left\langle x_i(t) - z_*, \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + f_i(x_i) \right\rangle, \end{aligned} \quad (47)$$

where  $\mathcal{I}(t)$  denotes the index set which contains all the nodes reaching the maximum for  $V(t)$ .

Let  $m \in \mathcal{I}(t)$ . Denote

$$Z_t = \{z : |z - z_*| \leq \sqrt{V(t)}\}$$

as the disk centered at  $z_*$  with radius  $\sqrt{V(t)}$ . Take  $y = x_m(t) + (x_m(t) - z_*)$ . Then from some simple Euclidean geometry it is obvious to see that  $P_{Z_t}(y) = x_m(t)$ , where  $P_{Z_t}$  is the projection operator onto  $Z_t$ . Thus, for all  $j \in \mathcal{N}_m(\sigma(t))$ , we obtain

$$\begin{aligned} \langle x_m(t) - z_*, x_j(t) - x_m(t) \rangle &= \langle y - x_m(t), x_j(t) - x_m(t) \rangle \\ &= \langle y - P_{Z_t}(y), x_j(t) - P_{Z_t}(y) \rangle \\ &\leq 0 \end{aligned} \quad (48)$$

according to inequality (i) in Lemma 3 since  $x_j(t) \in Z_t$ . On the other hand, based on inequality (i) in Lemma 4, we also have

$$\langle x_m(t) - z_*, f_m(x_m(t)) \rangle \leq F_m(z_*) - F_m(x_m(t)) \leq 0 \quad (49)$$

in light of the definition of  $z_*$ .

With (47), (48) and (49), we conclude that

$$D^+V(t) = \max_{i \in \mathcal{I}(t)} 2 \langle x_i(t) - z_*, \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + f_i(x_i) \rangle \leq 0, \quad (50)$$

which completes the proof.  $\square$

## D Proof of Lemma 9

In order to prove the desired conclusion, we just need to show  $\liminf_{t \rightarrow \infty} V_i(t) = d_*^2$  for all  $i$ . With Lemma 8, we conclude that  $\forall \varepsilon > 0, \exists M(\varepsilon) > 0$ , s.t.,

$$\sqrt{V_i(t)} \leq d_* + \varepsilon \quad (51)$$

for all  $i$  and  $t \geq M$ .

*Claim.* For all  $t \geq M$  and all  $i, j \in V$ , we have

$$\langle x_i(t) - z_*, x_j(t) - x_i(t) \rangle \leq -V_i(t) + (d_* + \varepsilon)\sqrt{V_i(t)}. \quad (52)$$

If  $x_i(t) = z_*$  (52) follows trivially from (51). Otherwise we take  $y_* = z_* + (d_* + \varepsilon)\frac{x_i(t) - z_*}{|x_i(t) - z_*|}$  and  $B_t = \{z : |z - z_*| \leq d_* + \varepsilon\}$ . Here  $B_t$  is the disk centered at  $z_*$  with radius  $d_* + \varepsilon$ , and  $y_*$  is a point within the boundary of  $B_t$  and falls the same line with  $z_*$  and  $x_{i_0}(t)$ . Take also  $q_* = y_* + x_i(t) - z_*$ . Then we have

$$\langle x_i(t) - z_*, x_j(t) - y_* \rangle = \langle q_* - y_*, x_j(t) - y_* \rangle = \langle q_* - P_{B_t}(q_*), x_j(t) - P_{B_t}(q_*) \rangle \leq 0 \quad (53)$$

according to inequality (i) in Lemma 3, which leads to

$$\begin{aligned} \langle x_i(t) - z_*, x_j(t) - x_i(t) \rangle &= \langle x_i(t) - z_*, x_j(t) - y_* \rangle + \langle x_i(t) - z_*, y_* - x_i(t) \rangle \\ &\leq \langle x_i(t) - z_*, y_* - x_i(t) \rangle \\ &= -V_i(t) + (d_* + \varepsilon)\sqrt{V_i(t)}. \end{aligned} \quad (54)$$

This proves the claim.

Now suppose there exists  $i_0 \in V$  with  $\liminf_{t \rightarrow \infty} V_i(t) = \theta_{i_0}^2 < d_*^2$ . Then we can find a time sequence  $\{t_k\}_1^\infty$  with  $\lim_{k \rightarrow \infty} t_k = \infty$  such that

$$\sqrt{V_{i_0}(t_k)} \leq \frac{\theta_{i_0} + d_*}{2}. \quad (55)$$

We divide the rest of the proof into three steps.

*Step 1.* Take  $t_{k_0} > M$ . We bound  $V_{i_0}(t)$  in this step.

With (52) and inequality (i) in Lemma 4, we see that

$$\begin{aligned}
 \frac{d}{dt}V_{i_0}(t) &= 2\left\langle x_{i_0}(t) - z_*, \sum_{j \in \mathcal{N}_{i_0}(\sigma(t))} a_{i_0j}(t)(x_j - x_{i_0}) + f_{i_0}(x_{i_0}(t)) \right\rangle \\
 &\leq 2 \sum_{j \in \mathcal{N}_{i_0}(\sigma(t))} a_{i_0j}(t) \left\langle x_{i_0}(t) - z_*, x_j(t) - x_{i_0}(t) \right\rangle + F_{i_0}(z_*) - F_{i_0}(x_{i_0}(t)) \\
 &\leq 2(N-1)a^* \left( -V_{i_0}(t) + (d_* + \varepsilon)\sqrt{V_{i_0}(t)} \right),
 \end{aligned} \tag{56}$$

for all  $t \geq t_{k_0}$ , which implies

$$\frac{d}{dt}\sqrt{V_{i_0}(t)} \leq -(N-1)a^* \left( \sqrt{V_{i_0}(t)} - (d_* + \varepsilon) \right), \quad t \geq t_{k_0}. \tag{57}$$

In light of Grönwall's inequality, (55) and (57) yield

$$\begin{aligned}
 \sqrt{V_{i_0}(t)} &\leq e^{-(N-1)^2 a^* T_D} \sqrt{V_{i_0}(t_{k_0})} + \left(1 - e^{-(N-1)^2 a^* T_D}\right)(d_* + \varepsilon) \\
 &\leq \frac{e^{-(N-1)^2 a^* T_D}}{2} \theta_{i_0} + \left(1 - \frac{e^{-(N-1)^2 a^* T_D}}{2}\right)(d_* + \varepsilon) \\
 &\doteq \Lambda_*.
 \end{aligned} \tag{58}$$

for all  $t \in [t_{k_0}, t_{k_0} + (N-1)T_D]$  with  $T_D = T + \tau_D$ , where  $T$  comes from the definition of uniformly jointly strongly connected graphs and  $\tau_D$  represents the dwell time.

*Step 2.* Since the graph is uniformly jointly strongly connected, we can find an instant  $\hat{t} \in [t_{k_0}, t_{k_0} + T]$  and another node  $i_1 \in V$  such that  $(i_0, i_1) \in G_{\sigma(t)}$  for  $t \in [\hat{t}, \hat{t} + \tau_D]$ . In this step, we continue to bound  $V_{i_1}(t)$ .

Similar to (52), for all  $t \geq M$  and all  $i, j \in V$ , we also have

$$\langle x_i(t) - z_*, x_j(t) - x_i(t) \rangle \leq -\sqrt{V_i(t)} \left( \sqrt{V_i(t)} - \sqrt{V_j(t)} \right) \tag{59}$$

when  $V_j(t) \leq V_i(t)$ . Then based on (52), (58), and (59), we obtain

$$\begin{aligned}
 \frac{d}{dt}V_{i_1}(t) &\leq 2 \sum_{j \in \mathcal{N}_{i_1}(\sigma(t))} a_{i_1j}(t) \left\langle x_{i_1}(t) - z_*, x_j(t) - x_{i_1}(t) \right\rangle \\
 &= 2 \sum_{j \in \mathcal{N}_{i_1}(\sigma(t)) \setminus \{i_0\}} a_{i_1j}(t) \left\langle x_{i_1}(t) - z_*, x_j(t) - x_{i_1}(t) \right\rangle + 2a_{i_1i_0}(t) \left\langle x_{i_1}(t) - z_*, x_{i_0}(t) - x_{i_1}(t) \right\rangle \\
 &\leq 2(N-2)a^* \left( -V_{i_1}(t) + (d_* + \varepsilon)\sqrt{V_{i_1}(t)} \right) - 2a_* \sqrt{V_{i_1}(t)} \left( \sqrt{V_{i_1}(t)} - \sqrt{V_{i_0}(t)} \right) \\
 &\leq -2 \left( (N-2)a^* + a_* \right) V_{i_1}(t) + 2\sqrt{V_{i_1}(t)} \left( (N-2)a^*(d_* + \varepsilon) + \Lambda_* a_* \right)
 \end{aligned} \tag{60}$$

for  $t \in [\hat{t}, \hat{t} + \tau_D]$ , where without loss of generality we assume  $V_{i_1}(t) \geq V_{i_0}(t)$  during all  $t \in [\hat{t}, \hat{t} + \tau_D]$ .

Then (60) gives

$$\frac{d}{dt}\sqrt{V_{i_1}(t)} \leq - \left( (N-2)a^* + a_* \right) \sqrt{V_{i_1}(t)} + \left( (N-2)a^*(d_* + \varepsilon) + \Lambda_* a_* \right), \quad t \in [\hat{t}, \hat{t} + \tau_D] \tag{61}$$

which yields

$$\begin{aligned}
 \sqrt{V_{i_1}(\hat{t} + \tau_D)} &\leq e^{-((N-2)a^* + a_*)\tau_D} (d_* + \varepsilon) + \left(1 - e^{-((N-2)a^* + a_*)\tau_D}\right) \frac{(N-2)a^*(d_* + \varepsilon) + \Lambda_* a_*}{(N-2)a^* + a_*} \\
 &= \frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*} \times \frac{e^{-(N-1)^2 a^* T_D}}{2} \theta_{i_0} \\
 &\quad + \left(1 - \frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*}\right) \times \frac{e^{-(N-1)^2 a^* T_D}}{2} (d_* + \varepsilon)
 \end{aligned} \tag{62}$$

again by Grönwall's inequality and some simple algebra.

Next, applying the estimate of node  $i_0$  in Step 1 on  $i_1$  during time interval  $[\hat{t} + \tau_D, t_{k_0} + (N-1)T_D]$ , we arrive at

$$\begin{aligned} \sqrt{V_{i_1}(t)} &\leq \frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*} \times \frac{e^{-2(N-1)^2 a^* T_D}}{2} \theta_{i_0} \\ &\quad + \left(1 - \frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*} \times \frac{e^{-2(N-1)^2 a^* T_D}}{2}\right) (d_* + \varepsilon) \end{aligned} \quad (63)$$

for all  $t \in [t_{k_0} + T_D, t_{k_0} + (N-1)T_D]$ .

*Step 3.* Noticing that the graph is uniformly jointly strongly connected, the analysis of steps 1 and 2 can be repeatedly applied to nodes  $i_3, \dots, i_{N-1}$ , and eventually we have that for all  $i_0, \dots, i_{N-1}$ ,

$$\begin{aligned} \sqrt{V_{i_m}(t_{k_0} + (N-1)T_D)} &\leq \left(\frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*}\right)^{N-2} \times \frac{e^{-(N-1)^3 a^* T_D}}{2} \theta_{i_0} \\ &\quad + \left(1 - \left(\frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*}\right)^{N-2} \times \frac{e^{-(N-1)^3 a^* T_D}}{2}\right) (d_* + \varepsilon) \\ &< d_* \end{aligned} \quad (64)$$

for sufficiently small  $\varepsilon$  because  $\theta_{i_0} < d_*$  and

$$\left(\frac{a_*(1 - e^{-((N-2)a^* + a_*)\tau_D})}{(N-2)a^* + a_*}\right)^{N-2} \times \frac{e^{-(N-1)^3 a^* T_D}}{2} < 1$$

is a constant. This immediately leads to that

$$V(t_{k_0} + (N-1)T_D) < d_*^2, \quad (65)$$

which contradicts the definition of  $d_*$ .

This completes the proof.  $\square$

## E Proof of Lemma 10

With Lemma 9, we have that  $\lim_{t \rightarrow \infty} V_i(t) = d_*^2$  for all  $i \in V$ . Thus,  $\forall \varepsilon > 0, \exists M(\varepsilon) > 0$ , s.t.,

$$d_* \leq \sqrt{V_i(t)} \leq d_* + \varepsilon \quad (66)$$

for all  $i$  and  $t \geq M$ . If  $d_* = 0$ , the desired conclusion follows straightforwardly. Now we suppose  $d_* > 0$ .

Assume that there exists a node  $i_0$  satisfying  $\limsup_{t \rightarrow \infty} |x_{i_0}(t)|_{\arg \min F_{i_0}} > 0$ . Then we can find a time sequence  $\{t_k\}_{k=1}^\infty$  with  $\lim_{k \rightarrow \infty} t_k = \infty$  and a constant  $\delta$  such that

$$|x_{i_0}(t_k)|_{\arg \min F_{i_0}} \geq \delta, \quad k = 1, \dots \quad (67)$$

Denote also  $B_1 \doteq \{z : |z - z_*| \leq d_* + 1\}$  and  $G_1 = \max\{|f_{i_0}(y)| : y \in B_1\}$ . Assumption A1 ensures that  $G_1$  is a finite number since  $B_1$  is compact. By taking  $\varepsilon = 1$  in (66), we see that  $x_i(t) \in B_1$  for all  $i$  and  $t \geq M(1)$ . As a result, we have

$$\left| \frac{d}{dt} x_{i_0}(t) \right| = \left| \sum_{j \in \mathcal{N}_{i_0}(\sigma(t))} a_{i_0 j}(t)(x_j - x_{i_0}) + f_{i_0}(x_{i_0}) \right| \leq 2(n-1)a^*(d_* + 1) + G_1. \quad (68)$$

Combining (67) and (68), we conclude that

$$|x_{i_0}(t)|_{\arg \min F_{i_0}} \geq \frac{\delta}{2}, \quad t \in [t_k, t_k + \tau], \quad (69)$$

for all  $k = 1, \dots$ , where by definition  $\tau = \frac{\delta}{2(2(n-1)a^*(d_*+1)+G_1)}$ .

Now we introduce

$$D_\delta \doteq \min \left\{ F_{i_0}(y) - F_{i_0}(z_*) : |y|_{\arg \min F_{i_0}} \geq \frac{\delta}{2} \text{ and } y \in B_1 \right\}.$$

Then we know  $D_\delta > 0$  again by the continuity of  $F_{i_0}$ . According to (56), (66), and (69), we obtain

$$\begin{aligned} \frac{d}{dt}V_{i_0}(t) &\leq 2(N-1)a^* \left( -V_{i_0}(t) + (d_* + \varepsilon)\sqrt{V_{i_0}(t)} \right) + F_{i_0}(z_*) - F_{i_0}(x_{i_0}(t)) \\ &\leq 2(N-1)a^*(2d_* + \varepsilon)\varepsilon - D_\delta, \end{aligned} \quad (70)$$

for  $t \in [t_k, t_k + \tau]$ ,  $k = 1, \dots$ . This leads to

$$\begin{aligned} V_{i_0}(t_k + \tau) &\leq V_{i_0}(t_k) + \left( 2(N-1)a^*(2d_* + \varepsilon)\varepsilon - D_\delta \right) \tau \\ &\leq (d_* + \varepsilon)^2 + \left( 2(N-1)a^*(2d_* + \varepsilon)\varepsilon - D_\delta \right) \tau \\ &< d_*^2 \end{aligned} \quad (71)$$

as long as  $\varepsilon$  is chosen sufficiently small. We see that (71) contradicts (66). The desired conclusion thus follows.  $\square$

## References

- [1] J. AUBIN AND A. CELLINA. *Differential Inclusions*. Berlin: Springer-Verlag, 1984.
- [2] V. N. BELYKH, I. V. BELYKH, AND M. HASLER, “Connection graph stability method for synchronized coupled chaotic systems,” *Physica D.*, vol. 195, pp. 159-187, 2004.
- [3] M. CAO, A. S. MORSE AND B. D. O. ANDERSON, “Reaching a consensus in a dynamically changing environment: a graphical approach,” *SIAM J. Control Optim.*, vol. 47, no. 2, 575-600, 2008.
- [4] F. CLARKE, YU. S. LEDYAEV, R. STERN, AND P. WOLENSKI. *Nonsmooth Analysis and Control Theory*. Berlin: Springer-Verlag, 1998.
- [5] J. CORTÉS, “Discontinuous dynamical systems—A tutorial on solutions, nonsmooth analysis, and stability,” *IEEE Control Systems Magazine*, vol. 28, no. 3, 36-73, 2008.
- [6] J. DANSKIN, “The theory of max-min, with applications,” *SIAM J. Appl. Math.*, vol. 14, 641-664, 1966.
- [7] F. DÖRFLER AND F. BULLO, “Synchronization and transient stability in power networks and non-uniform kuramoto oscillators,” *SIAM J. Contr. Opt.*, vol. 50, no. 3, pp. 1616-1642, 2012.
- [8] F. DÖRFLER AND F. BULLO, “Synchronization in complex networks of phase oscillators: A survey,” *Automatica*, vol. 50, no. 6, pp. 1539-1564, 2014.

- [9] H.-B. DÜRR AND C. EBENBAUER, “A smooth vector field for saddle point problems,” *The 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 4654-4660, 2011.
- [10] J. FAX AND R. MURRAY, “Information flow and cooperative control of vehicle formations,” *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1465-1476, 2004.
- [11] B. GHARESIFARD, J. CORTÉS, “Distributed continuous-time convex optimization on weight-balanced digraphs,” *IEEE Trans. Automatic Control*, vol. 59, no. 3, pp. 781-786, 2014.
- [12] C. GODSIL AND G. ROYLE. *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.
- [13] A. JADBABAIE, J. LIN, AND A. S. MORSE, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988-1001, 2003.
- [14] D. JAKOVETIĆ, J. XAVIER AND J. M. F. MOURA, “Cooperative convex optimization in networked systems: augmented Lagrangian algorithms with directed gossip communication,” *IEEE Trans. Signal Processing*, vol. 59, no. 8, pp. 3889-3902, 2011.
- [15] S. KAR AND J. M. F. MOURA, “Consensus + innovations distributed inference over networks: cooperation and sensing in networked systems,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 99-109, May 2013.
- [16] Y. KURAMOTO, “Self-entrainment of a population of coupled non-linear oscillators,” In H. Araki (Ed.), *Lecture notes in physics: vol. 39. Int. symposium on mathematical problems in theoretical physics* (pp. 420-422). Springer, 1975.
- [17] Z. LI, Z. DUAN, G. CHEN AND L. HUANG, “Consensus of multi-agent systems and synchronization of complex networks: A unified viewpoint,” *IEEE Trans. Circ. Syst.-I Regular papers*, vol. 57, no. 1, pp. 213-224, 2010.
- [18] Z. LIN, B. FRANCIS, AND M. MAGGIORE, “State agreement for continuous-time coupled nonlinear systems,” *SIAM J. Control Optim.*, vol. 46, no. 1, pp. 288-307, 2007.
- [19] H. LIU, M. CAO, AND C. W. WU, “Coupling strength allocation for synchronization in complex networks using spectral graph theory,” *IEEE Transactions on Circuits and Systems-I*, 61(5): 1520-1530. 2014.
- [20] S. MARTINEZ, J. CORTES, AND F. BULLO, “Motion coordination with distributed information,” *IEEE Control Syst. Mag.*, vol. 27, no. 4, pp. 75-88, 2007.
- [21] L. MOREAU, “Stability of multiagent systems with time-dependent communication links,” *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169-182, 2005.
- [22] A. NEDIĆ AND D.P. BERTSEKAS, “Incremental subgradient methods for nondifferentiable optimization,” *SIAM J. on Optimization*, vol. 12, pp. 109-138, 2001.
- [23] A. NEDIĆ AND A. OZDAGLAR, “Distributed subgradient methods for multi-agent optimization,” *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48-61, 2009.



- [24] A. NEDIĆ, A. OZDAGLAR AND P. A. PARRILO, “Constrained consensus and optimization in multi-agent networks,” *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922-938, 2010.
- [25] R. OLFATI-SABER, “Flocking for multi-agent dynamic systems: algorithms and theory,” *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401-420, 2006.
- [26] L. M. PECORA AND T. L. CARROLL, “Master stability functions for synchronized coupled systems,” *Physical Review Letters*, vol. 80, pp. 2019-2112, 1998.
- [27] A. Y. POGROMSKY AND H. NIJMEIJER, “Cooperative oscillatory behavior of mutually coupled dynamical systems,” *IEEE Trans. Circ. and Syst.-I Fundamental theory and applications*, vol. 48, no. 2, pp. 152-162, 2001.
- [28] A. Y. POGROMSKY, G. SANTOBONI, AND H. NIJMEIJER, “Partial synchronization: from symmetry towards stability,” *Physica D*, vol. 172, pp. 65-87, 2002.
- [29] S. S. RAM, A. NEDIĆ, AND V. V. VEERAVALLI, “Incremental stochastic subgradient algorithms for convex optimization,” *SIAM Journal on Optimization*, vol. 20, no. 2, pp. 691-717, 2009.
- [30] W. REN AND R. BEARD, “Consensus seeking in multi-agent systems under dynamically changing interaction topologies,” *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655-661, 2005.
- [31] R. T. ROCKAFELLAR. *Convex Analysis*. New Jersey: Princeton University Press, 1972.
- [32] N. ROUCHE, P. HABETS, AND M. LALOY. *Stability Theory by Liapunov’s Direct Method*. New York: Springer-Verlag, 1977.
- [33] L. SCARDOVI AND R. SEPULCHRE, “Synchronization in networks of identical linear systems,” *Automatica*, vol. 45, no. 11, pp. 2557-2562, 2009.
- [34] G. SHI AND Y. HONG, “Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies,” *Automatica*, vol. 45, no. 5, pp. 1165-1175, 2009.
- [35] G. SHI, A. PROUTIERE, AND K. H. JOHANSSON, “Continuous-time distributed optimization of homogenous dynamics,” *The 51st Annual Allerton Conference on Communication, Control, and Computing*, pp. 520-527, Oct. 2013.
- [36] G. SHI AND K. H. JOHANSSON, “Robust consensus for continuous-time multi-agent dynamics,” *SIAM J. Control Optim.*, vol. 51, no. 5, pp. 3673-3691, 2013.
- [37] G. SHI, K. H. JOHANSSON AND Y. HONG, “Reaching an optimal consensus: dynamical systems that compute intersections of convex sets,” *IEEE Transactions on Automatic Control*, vol. 58, no.3, pp. 610-622, 2013.
- [38] M. V. SOLODOV, “Incremental gradient algorithms with stepsizes bounded away from zero,” *Computational Optimization and Applications*, vol. 11, pp. 23-35, 1998.

- [39] Y. SU AND J. HUANG, “Stability of a class of linear switching systems with applications to two consensus problem,” *IEEE Trans. Aut. Contr.*, vol. 57, no. 6, pp. 1420-1430, 2012.
- [40] J. TSITSIKLIS, D. BERTSEKAS, AND M. ATHANS, “Distributed asynchronous deterministic and stochastic gradient optimization algorithms,” *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803-812, 1986.
- [41] T. VICSEK, A. CZIROK, E. B. JACOB, I. COHEN, AND O. SCHOCHET, “Novel type of phase transitions in a system of self-driven particles,” *Physical Review Letters*, vol. 75, pp. 1226–1229, 1995.
- [42] J. WANG AND N. ELIA, “Control approach to distributed optimization,” *The 48th Annual Allerton Conference*, pp. 557-561, 2010.
- [43] J. WANG AND N. ELIA, “A control perspective for centralized and distributed convex optimization,” in *IEEE Conf. on Decision and Control*, Orlando, Florida, pp. 3800-3805, 2011.
- [44] C. W. WU. *Synchronization in complex networks of nonlinear dynamical systems*. World Scientific, 2007.
- [45] C. W. WU AND L. O. CHUA, “Synchronization in an array of linearly coupled dynamical systems,” *IEEE Transactions on Circuits and Systems I*, vol. 42, pp. 430-447, 1995.
- [46] K. YOU AND L. XIE, “Network topology and communication data rate for consensusability of discrete-time multi-agent systems,” *IEEE Trans. Aut. Contr.*, vol. 56, no. 10, pp. 2262-2275, 2011.